Martín-Vázquez, R., Huertas-Tato, J., Aler, R., Galván, I.M. (2018) Studying the Effect of Measured Solar Power on Evolutionary Multi-objective Prediction Intervals. In *Intelligent Data Engineering and Automated Learning,* pp. 155-162, (11315).

DOI: https://doi.org/10.1007/978-3-030-03496-2_18

Studying the Effect of Measured Solar Power on Evolutionary Multi-objective Prediction Intervals

R. Martín-Vázquez, J. Huertas-Tato, R. Aler, and I. M. Galván(✉)

Computer Science Departament, Carlos III University of Madrid,
Avda. Universidad, 30, 28911 Leganés, Spain
igalvan@inf.uc3m.es

Abstract. While it is common to make point forecasts for solar energy generation, estimating the forecast uncertainty has received less attention. In this article, prediction intervals are computed within a multi-objective approach in order to obtain an optimal coverage/width trade-off. In particular, it is studied whether using measured power as an another input, additionally to the meteorological forecast variables, is able to improve the properties of prediction intervals for short time horizons (up to three hours). Results show that they tend to be narrower (i.e. less uncertain), and the ratio between coverage and width is larger. The method has shown to obtain intervals with better properties than baseline Quantile Regression.

Keywords: Solar energy · Prediction intervals
Multi-objective optimization

1 Introduction

In recent years, solar energy has shown a large increase in its presence in the electricity grid energy mix [1]. Having accurate point forecasts is important for solar energy penetration in the electricity markets and most of the research has focused on this topic. However, due to the high variability of solar radiation, it is also important to estimate the uncertainty around point forecasts. A convenient way of quantifying the variability of forecasts is by means of Prediction Intervals (PI) [2]. A PI is a pair of lower and upper bounds that contains future forecasts with a given probability (or reliability), named Prediction Interval Nominal Coverage (PINC). There are several methods for computing PI [3]: Delta method, Bayesian technique, Mean-Variance, and Bootstrap method. However, a recent evolutionary approach has shown better performance than the other methods in several domains [3], including renewable energy forecasting [4,5]. This approach, known as LUBE (Lower Upper Bound Estimation), uses artificial neural networks with two outputs, for the lower and upper bound of the interval, respectively. The network weights are optimized using evolutionary computation techniques such as Simulated Annealing (SA) [6] or Particle Swarm Optimization (PSO) [7].

The optimization of PI is inherently multi-objective, because of the trade-off between the two main properties of intervals: coverage and width. Coverage can be trivially increased by enlarging intervals, but obtaining high coverage with narrow intervals is a difficult optimization problem. Typical approaches to LUBE aggregate both goals so that optimization can be carried out by single-objective optimization methods (such as SA or PSO). In [8], it was proposed to use a multi-objective approach (using the Multi-objective Particle Swarm Optimization evolutionary algorithm or MOPSO [9]) for this purpose. The main advantage of this approach is that in a single run, it is able to obtain not one, but a set of solutions (the Pareto front) with the best trade-offs between coverage and width. If a solution with a particular PINC value is desired, it can be extracted from the front. In that work, due to the nature of the data, the aim was to estimate solar energy PI on a daily basis, and using a set of meteorological variables forecast for the next day.

In some works, meteorological forecasts are combined with measurements for training the models with the purpose of improving predictions [10,11]. In particular, in [12] it was observed that using measured power (in addition to meteo variables) was helpful to improve point forecasts for short term horizons. In the present work, we apply the MOPSO approach for estimating solar power PI, studying the influence of using measured solar power, in addition to meteorological forecasts, on the quality of intervals (coverage and interval width). In a similar way to point forecasts where using measured values improves the accuracy of the forecast if the prediction horizon is close to the measurement, we expect that using measured values will have an effect of reducing the uncer-tainty of prediction intervals. For that purpose, short prediction horizons of up to three hours will be considered in the experiments, in steps of 1 h. Linear Quantile Regression method [13] is also used as baseline with the purpose to comparison, using the R quantreg package [14].

The rest of the article is organized as follows: Sect. 2 describes the dataset used for experiments, Sect. 3 summarizes the evolutionary multi-objective app-roach for interval optimization, as well as the baseline method Quantile Regres-sion. Section 4 describes the experimental setup and the results. Conclusions are finally drawn in Sect. 5.

2 Data Description

The data used in this work is obtained from the Global Energy Forecasting Competition 2014 (GEFCom2014), specifically from task 15 of the probabilis-tic solar power forecasting problem [15]. The data provided included measured solar power and meteorological forecasts. Solar power was provided on an hourly basis, from 2012-04-01 01:00 to 2014-06-01 00:00 UTC (for training) and from 2014-06-01 01:00 to 2014-07-01 00:00 UTC for testing. The meteorological forecasts included 12 weather variables that had been obtained from the European Centre for Medium-range Weather Forecasts (ECMWF) [15]. Those variables were issued everyday at midnight UTC for each of the next 24 h. These 12 variables are: Total column liquid water (kg m-2), Total column ice water (kg m-2),

Surface pressure (Pa), Relative humidity at 1000 mbar (r), Total cloud cover (0'1), 10-metre U wind component (m s-1), 10-metre V wind component (m s-1), 2-metre temperature (K), Surface solar rad down (J m-2), Surface thermal rad down (J m-2), Top net solar rad (J m-2), and Total precipitation (m).

Data was provided for three power plants in Australia, although their exact location was not disclosed. In this article we are using station number 3 and the short term forecasting horizons (+1, +2, and +3 h). For some of the work carried out in this article, it is necessary to separate the available training data into training and validation. The validation set is used for model selection tasks, such as choosing the best neural network architecture or the best number of optimization iterations. In this work, the first 80% of the dataset has been used for training and the remaining 20% for validation.

3 Multi-objective Optimization for Prediction Intervals

The purpose of this section is to summarize the multi-objective evolutionary optimization of PI reported in [8]. This approach is based on LUBE [3], where a 3-layered neural network is used to estimate the lower and upper of bounds of prediction intervals for a particular input, but using a multi-objective evolutionary approach. The network receives as inputs meteorological variables. The outputs are the lower and upper bounds estimated by the network for those particular inputs. Although the observed irradiance for some particular inputs is available in the dataset, the upper and lower bounds are not. Hence, the standard back-propagation algorithm cannot be used to train the network (i.e. it is not a standard supervised regression problem, because the target output is not directly available). Therefore, in this approach, an evolutionary optimization algorithm is used to adjust the weights of the neural network by optimizing the two most relevant goals for PI: reliability and interval width. A prediction intervals is reliable if it achieves some specified reliability level or nominal coverage (i.e. PINC). This happens when irradiance observations lay inside the interval at least as frequently as the specified PINC. It is always possible to have high reliability by using very wide intervals. Therefore, the second goal used to evaluate PI is interval width (with the aim of obtaining narrow intervals). Reliability and interval width are formalized in the following paragraphs.

Let $M = \{(X_i, t_i)_{i=0\,..N}\}$ be a set of observations, where X_i is a vector with the input variables and t_i is the observed output variable. Let $PI_i = [Low_i, Upp_i)$ be the prediction interval for observation X_i (Low_i, Upp_i would be the outputs of the neural network). Then, the reliability (called Prediction Interval Coverage Probability or PICP) is computed by Eq. 1 and the Average Interval Width (AIW) by Eq. 2.

$$PICP = \frac{1}{N} \sum_{i=0}^{N} \chi_{PI_i}(X_i) \tag{1}$$

$$AIW = \frac{1}{N}\sum_{i=0}^{N}(Upp_i - Low_i) \tag{2}$$

where N is the number of samples, $\chi_{PI_i}(X_i)$ is the indicator function for interval PI_i (it is 1 if $t_i \in PI_i = [Low_i, Upp_i)$ and 0 otherwise). Upp_i and Low_i are the upper and lower bounds of the interval, respectively.

Given that there is a trade-off between reliability and width (PICP and AIW, respectively), the multi-objective approach (MOPSO) proposed in [8] is used to tackle the problem studied in this article. The MOPSO particles encode the weights of the networks, and the goals to be minimized are $1 - PICP$ (Eq. 1) and AIW (Eq. 2). In this work, the inputs to the networks are the meteorological variables given in the dataset (see Sect. 2) or any other information that may be useful for the estimation of solar irradiance (as solar power measurements).

The final result of MOPSO optimization is a non-dominated set of solutions (a Pareto front) as shown in Fig. 1. Each point (or solution) in the front represents the $x = AIW$ and $y = 1 - PICP$ of a particular neural network that achieves those values on the training dataset. If a particular target PINC is desired, then the closest solution in the Pareto front to that PINC is selected. That solution corresponds to a neural network that can be used on new data (e.g. test data) in order to compute PI for each of the instances in the test data. Figure 1 shows the solution that would be extracted from a Pareto front for PINC = 0.9.

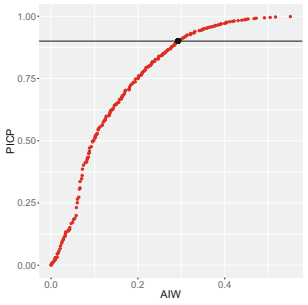

Fig. 1. Pareto front of solutions. Selected solution for PICP = 0.90.

In order to have a baseline to compare MOPSO results, Linear Quantile Regression (QR) has been used [13]. QR is a fast technique for estimating quantiles using linear models. While the standard method of least squares estimates the conditional mean of the response variable, quantile regression is able to estimate the median or other quantiles. This can be used for obtaining PI. Let q_1 and q_2 be the $\frac{1-PINC}{2}$ and $\frac{1+PINC}{2}$ quantiles, respectively. Quantile q_1 leaves a $\frac{1-PINC}{2}$ probability tail to the left of the distribution and quantile q_2 leaves $\frac{1-PINC}{2}$ probability tail to the right of the distribution. Therefore, the interval

$[q_1, q_2]$ has a coverage of PINC. Quantile Regression is used to fit two linear models that, given some particular input X_i, returns q_1 and q_2 with which the interval $[q_1, q_2]$ can be constructed.

4 Experimental Validation

As mentioned in the introduction, one of the goals of this article is to study the influence on the quality of intervals, of using measured solar power at the time of prediction $t_0 = 00:00$ UTC, in addition to the meteorological forecasts (that have already been described in Sect. 2). t_0 corresponds to 10:00 AM at the location in Australia, which is the time when meteo forecasts are issued everyday. With that purpose two derived datasets have been constructed, one with only the 12 meteorological variables and another one with those variables and the measured solar power at t_0. The latter (meteo + measured power) will be identified with $+P_{t_0}$. In both cases, the day of the year (from 1 to 365) has also been used as input, because knowing this information might be useful for computing the PI.

Table 1. Best combination of parameters for each prediction horizon

Horizon	MOPSO		MOPSO + P_{t_0}	
	Neurons	Iterations	Neurons	Iterations
1 h	15	8000	50	8000
2 h	8	8000	10	8000
3 h	6	8000	15	8000

We have followed a methodology similar to that of [8]. Different number of hidden neurons for the neural network (2, 4, 6, 8, 10, 15, 20, 30, and 50) and different number of iterations for PSO (4000, 6000, and 8000) has been tested. The process involves running PSO with the training dataset and using the validation dataset to select the best parameters (neurons and iterations). Given that PSO is stochastic, PSO has been run 5 times for each number of neurons and iterations, starting with different random number generator seeds. Similarly to [8], the measure used to select the best parameter combination has been the average hypervolume of the front on the validation set (the validation front is computed by evaluating each neural network from the training Pareto front, on the validation set). It is important to remark that, differently to [8], this has been done for each different prediction horizon. That means that this parameter optimization process has been carried out independently for each of three forecasting horizons considered in this work (+1 h, +2 h, +3 h). Table 1 displays the best combination of parameters for each horizon and whether P_{t_0} is used or not. It can be observed that the number of hidden neurons depends on the horizon and that the number of iterations is typically the maximum value

tried (8000). We have not extended the number of iterations for PSO because no further change was observed in the Pareto fronts by doing so.

In order to evaluate the experimental results, three target nominal coverage values (PINC) have been considered: 0.9, 0.95, and 0.99. The Quantile Regression approach must be run for each desired PINC value. The MOPSO approach needs to be run only once, because it provides a set of solutions (the Pareto front), out of which the solutions for particular PINCs can be extracted, as it has been explained in Sect. 3 (see Fig. 1).

Table 2. Evaluation measures on the test set for the four different approaches (QR, QR + P_{t_0}, MOPSO, MOPSO + P_{t_0}). Left: *Delta coverage*. Middle: Average interval width (AIW). Right: PICP/AIW ratio.

	Delta coverage			AIW			*PICP/AIW* ratio		
PINC	0.99	0.95	0.90	0.99	0.95	0.90	0.99	0.95	0.90
QR	0.027	0.072	0.089	0.756	0.611	0.495	1.291	1.438	1.640
QR + P_{t_0}	**0.020**	0.052	0.084	0.732	0.571	0.487	1.373	1.609	1.693
MOPSO	0.036	0.074	0.142	0.715	0.561	0.461	1.344	1.568	1.652
MOPSO + P_{t_0}	**0.020**	**0.051**	**0.061**	**0.646**	**0.495**	**0.427**	**1.530**	**1.861**	**2.018**

The performance of the solutions for each horizon, has been evaluated using three evaluation measures. The first one, named *delta coverage* in Table 2, measures how much the solution PICP fails to achieve the target PINC (on the test set). If the PICP fulfills the PINC ($PICP >= PINC$) then *delta coverage* is zero, otherwise it is computed as $PINC - PICP$ (in other words: *delta coverage* = $\max(0, PINC - PICP)$). The latter measure evaluates PINC fulfillment, but it only tells part of the performance because it is trivial to obtain small (or even zero) *delta coverage* by using very wide intervals. Thus, the second evaluation measure uses the ratio between PICP and the average interval width (AIW), which is calculated as $PICP/AIW$. Solutions that achieve high PICPs by means of large intervals will obtain low values for this ratio. Good solutions, with an appropriate tradeoff between PICP and width will obtain high values on this measure. Additionally, the average interval width (AIW) will also be shown.

Table 2 display the values of the three evaluation measures averaged over the 5 runs and the 3 horizons, for the three different values of PINC (0.99, 0.95, and 0.90). In Table 2 it can be seen that *delta coverage* (left) is larger than zero for all methods, which means that there are horizons for which PINC is not achieved. The best *delta coverage* values for all PINC values are obtained by MOPSO + P_{t_0} (this means that PICP is closer to the target PINC). It is also observed that the use of the measured solar power at 00:00 UTC helps MOPSO + P_{t_0} to obtain smaller *delta coverage*. Using P_{t_0} also helps QR in this regard. The same trend can be observed with respect to the AIW (see Table 2 middle) and the $PICP/AIW$ ratio (Table 2 right). Therefore, MOPSO + P_{t_0} obtains the best coverage, using the narrowest intervals, and reaching the best tradeoff between PICP and AIW.

Next, we will compare both approaches (MOPSO and QR) breaking down results by horizon. Table 3 shows the $PICP/AIW$ ratio and the AIW for all methods, horizons and PINC values. In the case of MOPSO and MOPSO + P_{t_0}, the average and standard deviation of the 5 runs are displayed. With respect to the ratio, it can be seen that using the P_{t_0} helps MOPSO for all horizons and target PINC values. In the QR case, it helps for the first horizon but not (in general) for the rest. The best performer for all horizons and target PINC is MOPSO + P_{t_0}, except for the second horizon and PINC = 0.99, where it is slightly worse than QR without P_{t_0}. The same trend can be observed for the AIW except for the third horizon and PINC = 0.90, where MOPSO and MOPSO + P_{t_0} are very similar.

Finally, for MOPSO, the improvement in the $PICP/AIW$ ratio by using P_{t_0} is larger for the first horizon than for the rest. For horizon 1, the improvement in ratio is 25%, 39%, and 40% for PINC values 0.99, 0.95, and 0.90, respectively. The reduction in AIW follows a similar behavior: 18%, 20%, and 18%, respectively. For the rest of horizons, there is also improvement, but smaller in size, and the larger the horizon, the smaller the improvement.

Table 3. Average and standard deviation of the $PICP/AIW$ ratio and AIW per prediction horizon (1 h, 2 h, 3 h). PINC values = 0.99, 0.95, 0.90.

Horizon	Method	*PICP/AIW ratio*			AIW		
		0.99	0.95	0.90	0.99	0.95	0.90
1	QR	1.258	1.416	1.561	0.742	0.589	0.491
1	QR + P_{t_0}	1.648	1.773	1.959	0.607	0.447	0.422
1	MOPSO	1.415 (0.085)	1.569 (0.114)	1.635 (0.098)	0.671 (0.048)	0.538 (0.040)	0.428 (0.011)
1	MOPSO + P_{t_0}	**1.762** (0.127)	**2.174** (0.117)	**2.294** (0.194)	**0.552** (0.055)	**0.430** (0.039)	**0.351** (0.031)
2	QR	**1.452**	1.481	1.761	**0.666**	0.585	0.473
2	QR + P_{t_0}	1.375	1.574	1.565	0.677	0.613	0.485
2	MOPSO	1.283 (0.099)	1.508 (0.114)	1.571 (0.044)	0.747 (0.058)	0.596 (0.051)	0.484 (0.026)
2	MOPSO + P_{t_0}	1.446 (0.108)	**1.734** (0.114)	**1.918** (0.184)	0.691 (0.057)	**0.515** (0.048)	**0.451** (0.062)
3	QR	1.162	1.417	1.599	0.861	0.659	0.521
3	QR + P_{t_0}	1.097	1.481	1.555	0.911	0.652	0.554
3	MOPSO	1.333 (0.057)	1.628 (0.106)	1.749 (0.235)	0.727 (0.031)	0.550 (0.027)	**0.470** (0.040)
3	MOPSO + P_{t_0}	**1.383** (0.092)	**1.674** (0.216)	**1.842** (0.209)	**0.696** (0.047)	**0.540** (0.065)	0.478 (0.026)

5 Conclusions

In this article, we have used a multi-objective approach, based on Particle Swarm Optimization, to obtain prediction intervals with an optimal tradeoff between interval width and reliability. In particular, the influence on short prediction horizons, of using measured solar power as an additional input, has been studied. This has shown to be beneficial, because prediction interval tend to be narrower (hence, less uncertainty on the forecast), and the ratio between coverage and width is larger. This is true for the three short prediction horizons studied, but the improvement is larger for the shortest one (+1 h). Results have been compared to Quantile Regression and shown to be better for all evaluation criteria.

While Quantile Regression also benefits from using measured solar radiation, this happens only for the 1 h horizon, but not for +2 or +3 h.

Acknowledgements. This work has been funded by the Spanish Ministry of Science under contract ENE2014-56126-C2-2-R (AOPRIN-SOL project).

References

1. Raza, M.Q., Nadarajah, M., Ekanayake, C.: On recent advances in pv output power forecast. Sol. Energy **136**, 125–144 (2016)
2. Pinson, P., Nielsen, H.A., Møller, J.K., Madsen, H., Kariniotakis, G.N.: Non-parametric probabilistic forecasts of wind power: required properties and evaluation. Wind. Energy **10**(6), 497–516 (2007)
3. Khosravi, A., Nahavandi, S., Creighton, D., Atiya, A.F.: Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Trans. Neural Netw. **22**(3), 337–346 (2011)
4. Wan, C., Xu, Z., Pinson, P.: Direct interval forecasting of wind power. IEEE Trans. Power Syst. **28**(4), 4877–4878 (2013)
5. Khosravi, A., Nahavandi, S.: Combined nonparametric prediction intervals for wind power generation. IEEE Trans. Sustain. Energy **4**(4), 849–856 (2013)
6. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. Science **220**(4598), 671–680 (1983)
7. Eberhart, R.C., Shi, Y., Kennedy, J.: Swarm Intelligence. Elsevier, Amsterdam (2001)
8. Galván, I.M., Valls, J.M., Cervantes, A., Aler, R.: Multi-objective evolutionary optimization of prediction intervals for solar energy forecasting with neural networks. Inf. Sci. **418**, 363–382 (2017)
9. Coello Coello, C.A., Lechuga, M.S.: MOPSO: a proposal for multiple objective particle swarm optimization. In: Proceedings of the 2002 Congress on Proceedings of the Evolutionary Computation on CEC 2002, vol. 2, pp. 1051–1056. IEEE Computer Society, Washington (2002)
10. Aguiar, L.M., Pereira, B., Lauret, P., Díaz, F., David, M.: Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. Renew. Energy **97**, 599–610 (2016)
11. Wolff, B., Kühnert, J., Lorenz, E., Kramer, O., Heinemann, D.: Comparing support vector regression for pv power forecasting to a physical modeling approach using measurement, numerical weather prediction, and cloud motion data. Sol. Energy **135**, 197–208 (2016)
12. Martín-Vázquez, R., Aler, R., Galván, I.M.: Wind energy forecasting at different time horizons with individual and global models. In: Iliadis, L., Maglogiannis, I., Plagianakos, V. (eds.) AIAI 2018. IAICT, vol. 519, pp. 240–248. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-92007-8 21
13. Koenker, R.: Quantile Regression. Econometric Society Monographs, vol. 38. Cambridge University Press, Cambridge (2005)
14. Koenker, R.: quantreg: Quantile Regression. R package version 5.36 (2018)
15. Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., Hyndman, R.J.: Probabilistic energy forecasting: global energy forecasting competition 2014 and beyond. Int. J. Forecast. **32**(3), 896–913 (2016)